Request to Archive
With The National Centers for Environmental Information
For Climate Forecast System Operational v2 (CFSv2)
Provided by NOAA/NWS/NCEP/CPC

2014-02-15

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Glenn Rutledge
NCDC/CSMD/DAAB
Team Lead
828-271-4097
glenn.rutledge@noaa.gov

2. Name the organization or group responsible for creating the dataset.

NWS NCEP EMC/CPC

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

The Climate Forecast System Operational v2 (CFSv2) system is a coupled atmospheric-ocean-land-sea-ice model that contains 6-hourly forecasts, and seasonal 45-day and 9-month output in GRIB format. The data is already described as part of the original Submission Agreement (SA) of 2009/10. This request is an extension of the at request. It is 0.5, 1.0 and 2.5 degree data on a global scale. These data stopped being ingested in March 2011. This is the point where the dataset transitions from the "CFS Reanalysis" to the CFS Operational" dataset.

The total static/backlog data volume as recommended for archiving is estimated at 468 TB, and operational data volume is approximately 64 TB per year. The output types covered by this request include:
- CFSv2 6-hourly PGB and FLX Reforecast (Jan 1999 - Mar 2011): 246 TB (static)
- CFSv2 Calibration Climatologies (1982/99 - 2010): 30 TB (static)
- CFSv2 Operational Forecasts Backlog (Apr 2011 - Mar 2014) 168 TB (static)
- CFSv2 Operational CDAS Backlog (Apr 2011 - Mar 2014): 24 TB (static)
- CFSv2 Operational Forecasts Ongoing (Mar 2014 - daily): 56 TB/yr
- CFSv2 Operational CDAS Ongoing (Mar 2014 - daily): 8 TB/yr

For the 1999-2011 CFS Reforecast product suite, NCEP generated 45-day and 90-day reforecasts with a period of record (PoR) beginning January 1, 1999 and extending through March 31, 2011. For each day of the PoR, 90-day reforecasts were produced at the 00Z cycle, and 45-day reforecasts were produced on the 06Z, 12Z, and 18Z cycles. 9-month reforecasts were generated with a PoR from January 1, 1982 through March 31, 2011. 9-month reforecasts were only produced every 5 days in the PoR, and on those days once for each 00Z, 06Z, 12Z, and 18Z cycle. Each 9-month reforecast consists of only 6 months of reforecast data. The dataset includes timeseries and monthly means files for the first 3 months of 2011. Timeseries and monthly means contain a full 9 months of reforecast data.

The monthly mean and time series calibration climatologies were performed on PBG, FLX, and OCN files. There are two datasets: one is a full PoR dataset (1982-2010), and the second is a dataset comprised of 1999-2010 data. Both

datasets end calculations with 2010 data. The 1999-2010 dataset was created due to the addition of AMSU satellite assimilation data. This latter dataset is more useful with calibrations in the tropics.

The operational forecast since 2011 contains four-times-daily, 9-month control runs, consisting of all 6-hourly PGB, OCN, FLX, and IPV files, and the monthly means and time-series files. The operational CDAS analysis since 2011 contains daily, 6-hourly IPV, OCN, PGRB, SFLUXGRB files; and once-a-month times-series and monthly means files. For the operational 6-hourly forecast data, each 9-month control run consists of only 6 months of forecast data due to loss of model skill at these projections, however, the timeseries and monthly means extend to 9 months. The PoR for the dataset begins on April 1, 2011 and continues onward. An estimated 192TB of both CFSv2 Operational forecast and CDAS data, from 1 April, 2011 to 1 March, 2014 will need to be transferred. The operational CFSv2 data starting in April 2011 to present has highest priority according to NCEP since this dataset is supposed to be purged of data older than two years and this purging is under a temporary suspension directive.

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)
From 1982-01-01
Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:
v2

6. Describe the level to which the data are processed. For example, are these unprocessed raw observations, derived parameters, quality controlled or inter-calibrated data, etc.?
These are highly processed model output using satellite, radar and surface observation data. Output is produced on native and gaussian grids, and written to WMO standard grib grids.

7. Approximate date when the dataset was or will be released to the public:
2011-03-01

8. Who are the expected users of the archived data? How will the archived data be used?
Researchers, improving model output thereby improving NWS skill for life and property. Hydrological community, the climate analysis community, private sector (commodities/trading/reinsurance-Wall street),

9. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?
NOAA Science Advisory Board, NWS TAFF's, NOAA Operational / Reanalysis Steering group

10. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?
Part of the data in this request is for a continuation of the CFS Reanalysis and Reforecast labeled as operational CFSv2 since it is an on-going product at NCEP.

11. List the input datasets and ancillary information used to produce the data.
See CFSv2 documentation on data assimilation.

12. List web pages and other links that provide information on the data.
Metadata follows WMO Standards in GRIB2 format

13. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. http://cfs.ncep.noaa.gov/cfsr/docs/
2. http://nomads.ncdc.noaa.gov/data.php?name=access#cfs
3. CFSRR SA between NCDC and NCEP

14. Indicate the data file format(s).

1. GRIB 2

15. Are the data files compressed?

No

16. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

will match the file format as described in CFSRR SA

17. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

Earlier CFSRR data are available via NOMADS

18. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume:	468TB
Number of Data Files:	500000

Continuous Data: data volume rate for a continuous data production.

Total Data Volume Rate:	64TB per Year
Data File Frequency:	100 per Day
Data Production Start:	2011-04-01

19. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

The data is real-time fcst product. The data exists from 1979 to present in several forms as the CFS Reanalysis. This request will be the continuation of that.

20. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location:	College Park, MD
System Name:	Zeus /HPSS
System Owner:	NCEP/EMC
Additional Information:	For Historical backlog data; Operaional data from NCEP/NCO

21. What are the possible methods for submitting the data to NCEI? Select all that apply.

1. FTP PUSH

22. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. User interface to order and stage data for download
2. Direct download links
3. Advanced web services (e.g., THREDDS Catalog Service)

23. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?
No known constraints apply to the data.

24. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.
This data is being heavily requested by the research and private sector community to improve both climate and weather forecasting for the U.S. It is also being requested to satisfy OAR's Research MAPP program- that funds the NOMADS/NCMP Projects. It has also been requested by the UofMD CIRUN program and the NOAA Climate Testbed.

25. Are the data archived at another facility or are there plans to do so? Please explain.
No

26. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?
Maybe. The "CFS Archive Recommendation" as produced by NCDC was never funded. Also, the joint NCDC-NCEP-CPO "CFS Needs Assessment Workshop" has recommended this to be archived.

27. Do you have a data management plan for your data?
No

28. Have funds been allocated to archive the data at NCEI?
The NCMP program will fund as much of these data as possible, given level (FY13) funding for FY14. Cost estimates are currently being generated in the NOMADS-CLASS pilot projects and are expected to be known by Dec 2013 timeframe.

29. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.
NOMADS and NCMP. Climate Program Office

30. Is there a desired deadline for NCEI to archive and provide access to the data?
Archive by: 2014-03-26
Accessible by:

31. Add any other pertinent information for this request.
None